



On the Reactivity of Think-Aloud as a Data Collection Tool and the Potential Modulating Effect of Think-Aloud Instruction on the Reactivity of this Tool

Yinghui Li¹ 
Hengbin Yan² 

¹School of English and Education/Bilingual Cognition and Development Lab, Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, China

Email: liyinhui@gdufs.edu.cn Tel: 86-20-36207201

²Faculty of English Language and Culture, Guangdong University of Foreign Studies Baiyun Avenue North 2#, Baiyun District, 510420, Guangzhou, China

Email: yhb@gdufs.edu.cn Tel: 86-20-36209276



(✉ Corresponding Author)

Abstract

To explore the reactivity of think-aloud (i.e., whether think-aloud as a data collection tool would influence the data it collects), the present study compared 60 Chinese-English bilinguals' performances in solving Advanced Progressive Matrices (APM) in two thinking modes, either thinking aloud or not. Furthermore, this study examined whether the potential task performance difference was modulated by the think-aloud instruction language, either first language (L1) or second language (L2). The participants were separated into two groups and instructed respectively in L1 and L2 to solve APM in two thinking modes (with the order of the two modes counterbalanced within group). Results show that there was no significant difference in task performance between the two thinking modes; moreover, the think-aloud instruction language did not have any significant modulating effect on the task performance difference between the two modes. Implications of this study for empirical research employing think-aloud as a data collection tool are discussed.

Keywords: Think-Aloud, Concurrent verbalization, Reactivity, Instruction, Bilingual, Advanced progressive matrices.

JEL Classification: Z13.

Citation | Yinghui Li; Hengbin Yan (2018). On the Reactivity of Think-Aloud as a Data Collection Tool and the Potential Modulating Effect of Think-Aloud Instruction on the Reactivity of this Tool. *International Journal of Social Sciences and English Literature*, 2(1): 27-32.

History:

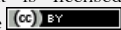
Received: 5 September 2018

Revised: 8 October 2018

Accepted: 13 November 2018

Published: 10 December 2018

Licensed: This work is licensed under a [Creative Commons](https://creativecommons.org/licenses/by/3.0/)

Attribution 3.0 License 

Publisher: Asian Online Journal Publishing Group

Contribution/Acknowledgement: Both authors contributed to the conception and design of the study. Many thanks to Prof. Qiufang Wen from Beijing Foreign Studies University, China, for her invaluable suggestion to the current research.

Funding: The current research was supported by three grants, including a grant (grant number: BCD201702) from Bilingual Cognition and Development Lab, Guangdong University of Foreign Studies, a grant (grant number: 290-XGS17023) directly from Guangdong University of Foreign Studies, and a grant (grant number: GD18YWW02) from Guangdong Planning Office of Philosophy and Social Science.

Competing Interests: The authors declare that they have no conflict of interests.

Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study was reported; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Ethical: This study follows all ethical practices during writing.

Contents

1. Introduction	28
2. Literature Review	28
3. The Present Study	29
4. Results	30
5. Discussion.....	30
6. Conclusion	31
References.....	31

1. Introduction

Intended to elucidate one's internal cognitive processes, think-aloud (also named concurrent verbalization) along with retrospective verbalization is one of the data collection tools widely used in studies on psychology (e.g., Ericsson and Simon, 1998; Eveland and Dunwoody, 2000; Aujla *et al.*, 2018) first/second language acquisition (Cohen, 1984;1986; Yi'an, 1998; Benson *et al.*, 2009) translation and interpreting (Ivanova, 2000; Bartłomiejczyk, 2006) and so on. A number of handbooks and manuals on think-aloud (Ericsson and Simon, 1984/1993; Ericsson, 2002; Guo, 2007) have thus emerged providing instructions on how to collect reliable and valid data by using this tool. The use of this tool, however, remains controversial mainly due to our insufficient knowledge on two issues. First, the issue of completeness and accuracy, that is, whether what one reports in a think-aloud process (that is, think-aloud protocols) is consistent with one's thinking process and product (Ericsson and Simon, 1984/1993; Stratman and Hamp-Lyons, 1994; Wilson, 1994). Second, the issue of reactivity, that is, whether one's task performance under the condition of thinking aloud differs from his or her performance in the same task under the condition of not thinking aloud, and what factors may modulate the potential difference (Smagorinsky, 1994a;1994b;2001; Bowles, 2008;2010; Sanz *et al.*, 2009). The present study focuses on the second issue, aiming to investigate how task performance may differ under different thinking modes (either thinking aloud or not thinking aloud), along with how the think-aloud instruction may modulate the potential difference in task performance.

2. Literature Review

2.1. Think-Aloud

Think-aloud involves a process in which participants report verbally their immediate thoughts and intermediate products of their continually-generated thoughts (Ericsson and Simon, 1984/1993). Major manuals on think-aloud (Ericsson and Simon, 1980;1984/1993; Ericsson, 2002) illustrate three levels at which think-aloud can be operated. Based on a series of empirical studies, these manuals concluded that (1) think-aloud at Level 1 (that is, simply articulating thoughts that are already encoded in words, where no intermediate processes of transforming non-verbal codes into verbal ones are needed) and at Level 2 (that is, describing one's thoughts and labeling the thoughts, of which some may not be encoded yet but is describable) can elicit reliable and valid verbal protocols; (2) think-aloud at Level 3 (that is, explaining one's thinking process through recoding thoughts, recalling thoughts reported before, and connecting what have been reported to what to be reported), however, may not be able to elicit reliable or valid verbal protocol. Accordingly, the present study focuses on think-aloud at the first two levels aforementioned, and only studies following the instructions eliciting verbal protocols at these two levels are thus to be reviewed.

Inconsistent empirical findings on the reactivity of think-aloud can be found between the studies indicating that think-aloud influenced one's thinking process and thus task performance (Stinessen, 1985; Schooler *et al.*, 1993; Kim, 2002; Rossomondo, 2007; Sachs and Polio, 2007; Sanz *et al.*, 2009) and the studies where no significant difference was found in task performance with or without think-aloud operated (Flaherty, 1975; Schweiger, 1983; Russo *et al.*, 1989; Leow and Morgan-Short, 2004; Bowles and Leow, 2005; Sachs and Suh, 2007; Bowles, 2008). After reviewing these studies, researchers (as reviewed below) propose that the inconsistency in the empirical findings mainly originate from various factors that can modulate the potential influence of think-aloud on task performance but were not well controlled in the previous studies, although these studies can be also improved in the research design, sample size and statistical analysis.

2.2. Think-Aloud Instruction

Both theoretical and empirical studies have explored potential factors that can modulate the influence of think-aloud on task performance, which include:

- (1) level of think-aloud (Ericsson and Simon, 1984/1993)
- (2) think-aloud instruction (Schweiger, 1983; Ericsson and Simon, 1984/1993; Guo, 2007)
- (3) when participants think aloud, what type of task they are to accomplish (Russo *et al.*, 1989)
- (4) what aspect of task performance is investigated, for example latency or accuracy (Sanz *et al.*, 2009)
- (5) participants' familiarity with and expertise in the task (Stratman and Hamp-Lyons, 1994; Smagorinsky, 1994b; Fallshore and Schooler, 1995)
- (6) participants' language proficiency (Stratman and Hamp-Lyons, 1994) and how consistent participants' perception of their language proficiency is with their actual language proficiency (Fallshore and Schooler, 1995)
- (7) participants' cultural background, in particular whether they feel comfortable with articulating while thinking (Kim, 2002)
- (8) participants' thinking habit (thinking in images or not: (Brandimonte *et al.*, 1992; Fallshore and Schooler, 1995) thinking in words or thinking in physical movements: Flaherty, 1975; thinking in a way of intuition/insight or not: (Schweiger, 1983)
- (9) think-aloud instructors' personal traits (Stratman and Hamp-Lyons, 1994; Smagorinsky, 1994b).

Among these modulating factors, the current study focuses on the think-aloud instruction and, in particular, the language used in the instruction. As the think-aloud instruction is found to play a critical role in eliciting reliable and valid verbal protocols (Schweiger, 1983; Ericsson and Simon, 1984/1993; Guo, 2007) the language employed in the instruction, which is one of the important elements of instruction, may also theoretically modulate the potential influence of think-aloud on task performance. So far no research has addressed the issue, which prompts the need for the present study to empirically examine and compare the potential influence of think-aloud on task performance when the think-aloud instruction language differs. The issue is particularly worth exploring to studies of second language acquisition and studies of translation and interpreting, in which bilingual or multilingual participants are usually involved. The justification is that in these studies where think-aloud is one of the commonly-used data collection tools, if the think-aloud instruction language, either first language (L1) or second/foreign language (L2), is not found to modulate the potential influence of think-aloud on task performance,

it would be better justified for instructors to use L2 in their think-aloud instructions (instead of solely L1) so that participants' verbal thinking in L2 if any (especially the mental activity generated in a bilingual or L2 task) can be elicited more efficiently compared with when participants are instructed only in L1 (which is common in think-aloud instructions).

3. The Present Study

This study is intended to investigate the reactivity of think-aloud and the potential modulating effect of the think-aloud instruction language on the reactivity. In order to control the other potential modulating factors (as reviewed above) in which the current study is not interested, (1) we follow strictly the instruction of eliciting think-aloud at Level 1 and 2 as introduced in the major manuals of think-aloud (Ericsson and Simon, 1984/1993; Guo, 2007) (2) all participants are invited to accomplish the same task and rated in accuracy; (3) L2 (in this case, English) proficiency is assumed homogeneous (at the level no lower than intermediate) among participants, given that all of our participants have passed a competitive English college entrance exam in mainland China and have been enrolled in a national key university of foreign studies. We further prevent language proficiency from intervening our results by ① choosing Advanced Progressive Matrices (APM for short hereafter) (Raven *et al.*, 1978) which is an established non-verbal problem-solving task, as the major task in the current study and by ② allowing participants to use any language they feel comfortable with in think-aloud (no matter what language is used in the instruction). Also to control the other potential modulating factors, (4) a practice session of APM is available so that participants' task familiarity is not likely to intervene in the results; (5) since all of our participants have been brought up and educated mostly in China and all major in English in the same university, their cultural background and thinking mode are assumed to be homogenous in general; (6) all participants are instructed by the same think-aloud instructor under the same procedure.

The present study aims to answer two research questions as below:

- (1) Does Chinese English-as-a-foreign-language learners' performance in APM differ between different thinking modes (either thinking aloud or not thinking aloud)?
- (2) How much does the language used in the think-aloud instruction (either L1 or L2) modulate the potential difference in task performance under different thinking modes?

3.1. Design

The within-subjects independent variable was thinking mode (thinking aloud or not thinking aloud) and the between-subjects independent variable was instruction language (L1 or L2). The dependent variable was accuracy in APM tests (as manifested in APM test score).

3.2. Participants

63 female first-year university students majoring in English at a national key university of foreign studies were recruited on a voluntary basis to participate in the experiment. They were all Chinese-English unbalanced bilinguals who learned English as a foreign language mainly in China for at least 10 years. Data from three participants who did not follow the instruction of thinking aloud or the instruction of not thinking aloud were excluded, resulting in data of 60 participants analyzed.

3.3. Materials

Major materials included (1) a questionnaire about participants' demographic details (mainly age, general education background, languages and self-assessed proficiency of these languages, age of first exposure to English, major areas of learning and using English, and their familiarity with APM) and (2) two APM tests respectively used in the condition of thinking aloud (Test Think-aloud) and in the condition of not thinking aloud (Test No Think-aloud). The two APM tests used in the current study, each of which consisted of 10 items, were adopted from the original APM test and were proved homogenous in a pilot study with 16 participants from the same population as the aforementioned participants.

3.4. Apparatus and Procedure

The participants were separated into two groups (respectively labeled as Group L1 and Group L2) and instructed in L1 (in this case, Chinese) and L2 (in this case, English) respectively. Each group was asked to finish both Test Think-aloud and Test No Think-aloud, with the order of the two tests counterbalanced within group. Between the two tests was a 10-minute interval for the participants to take a rest. In Test Think-aloud, participants first watched a video demo of think-aloud, followed by an introduction of think-aloud principles, which mainly included (1) only reporting thoughts that have indeed gone through their mind; (2) not fabricating any idea for the report or explaining any of their ideas (for example, explaining the rationale or justification for any idea or the logic underlying different ideas); (3) using or shifting between whatever language(s) they felt comfortable with, no matter what language was employed in the instruction. Afterwards, as a practice the participants solved three APM items while thinking aloud, during which they were allowed to pause and ask the instructor questions if needed. The whole think-aloud instruction (either in L1 or in L2) followed guidelines offered by Ericsson and Simon (1984/1993) and Guo (2007) and had been modified based on the feedback about our think-aloud instruction provided by three participants (from the same population as the participants in the major study) in an interview after a pilot study.

After the think-aloud instruction and practice, the participants took Test Think-aloud independently. During both Test Think-aloud and Test No Think-aloud, which were held in a language laboratory, the think-aloud instructor paid close attention to the participants' verbalization via a monitoring system so that the participant who remained silent for longer than 30 seconds in Test Think-aloud would be prompted to report her thoughts as soon as possible; on the other hand, the participant who did not keep silent in Test No Think-aloud would also be

reminded (unless when they asked the instructor questions). All of the participants had consented to the use of the monitoring and recording system in the two tests.

4. Results

4.1. Thinking Mode

As shown in Table 1, no significant difference was found in our participants' APM test score (in total 10 points) between Test Think-aloud and Test No Think-aloud ($t(59) = -0.72, p = .47 > .05$), although the performance in the former (Mean = 7.63, Standard deviation = 1.35) test was slightly better than that in the latter (Mean = 7.78, Standard deviation = 1.33). The results suggest that no reactivity of think-aloud was found when our participants reported verbally their thinking processes while solving APM.

Table-1. APM performances under two thinking modes (thinking aloud or not) (N=60)

Thinking mode	Mean	Standard deviation	<i>t</i>	<i>p</i>
not thinking aloud	7.63	1.35	-0.72	0.47
thinking aloud	7.78	1.33		

4.2. Instruction Language

As demonstrated in Table 2, the performance in APM was not found significantly different between the two thinking modes within Group L1 ($t(29) = -0.67, p = .51 > .05$), nor was within Group L2 ($t(29) = -0.37, p = .71 > .05$).

Table-2. APM performances under two thinking modes (thinking aloud or not) when the instruction language was either L1 or L2

Instruction language	N	Thinking mode	Mean	Standard deviation	<i>t</i>	<i>p</i>
L1	30	not thinking aloud	7.43	1.45	-0.67	0.51
		thinking aloud	7.63	1.38		
L2	30	not thinking aloud	7.80	1.30	-0.37	0.71
		thinking aloud	7.92	1.29		

To further investigate the potential interaction of thinking mode and instruction language, a two-way ANOVA was performed on the APM test score of these two independent variables. The observed covariance matrices of test scores were not found significantly different across groups (Box's M = .50, $p = .99 > .05$). With values from the corrected *F*-tests considered ($\chi^2 = .00$ in the Mauchly's test of sphericity), results (Table 3) showed that the two-way interaction between the variables was not significant ($F(1, 59) = .06, p = .80 > .05$). Moreover, results from the Levene's Test (Table 4) show that the error variance of the dependent variable (that is, APM test score) was equal across groups (in think-aloud condition, $F = .46, p = .50 > .05$; in no think-aloud condition, $F = .71, p = 0.40 > .05$). These results indicate that the thinking mode and the instruction language were not found to have any significant interaction on the participants' task performance in APM, suggesting that no modulating effect of the think-aloud instruction language was found on the relation between the thinking mode and the task performance in APM.

Table-3. The within-subjects effect of thinking mode (thinking aloud or not) and its interaction with instruction language (L1 or L2) (N = 60)

variable	<i>F</i>	<i>p</i>
Thinking mode	0.57	0.45
Thinking mode * Instruction language	0.06	0.80

Table-4. Levene's Test of Equality of Error Variances^a

Test score	<i>F</i>	Sig.
in think-aloud condition	0.46	0.50
in no think-aloud condition	0.71	0.40

a. Design: Intercept + instruction language
Within Subjects Design: thinking mode

5. Discussion

The present study aimed to explore the potential influence of think-aloud on one's task performance in APM and the potential modulating effect of the think-aloud instruction language on the aforementioned influence of think-aloud. Two major findings were attained. First, the performance of solving APM did not significantly differ between thinking aloud and not thinking aloud. As illustrated by Ericsson and Simon (1984/1993) think-aloud at Level 1 and Level 2 do not occupy the cognitive resources needed for accomplishing the major task, so one's task performance when one thinks aloud would not differ significantly from his or her task performance when he or she does not think aloud. Our finding is consistent with the general findings in Flaherty (1975); Russo *et al.* (1989) and Schweiger (1983) indicating that think-aloud as a data collection tool does not have significant reactivity problem. When Kim (2002) found difference in task performance between thinking aloud and not thinking aloud, the study attributed the difference to participants' Asian cultural background emphasizing the value of keeping silent in public, which led to a suggestion that think-aloud should not be used in empirical studies where participants came from Asian cultural background. Since the participants in our study were English learners in China, our findings at least cast doubt on such a suggestion.

Second, the results in our study did not show that think-aloud instruction languages had any significant modulating effect on the relation between the thinking mode and the performance in the major task, indicating that to future bilingual studies using think-aloud as a data collection tool, L2 can be one of the language options in the think-aloud instruction, in particular when researchers aim to explore bilinguals' verbal thinking in L2. Nevertheless, it is noteworthy that in order to preclude the potential intervention from language proficiency, we

suggest that participants should be allowed to use whatever language in thinking aloud and they should not report in their non-dominant language(s) until they spontaneously want to.

6. Conclusion

The present study is intended to examine whether APM task performance differs in different thinking modes (thinking aloud or not). Furthermore, it aims to investigate how much the think-aloud instruction language (L1 or L2) may modulate the potential task performance difference under the two thinking modes. The results show that no significant difference in APM task performance exists between the mode of thinking aloud and the mode of not thinking aloud. Besides, the results show that the instruction language does not have significant modulating effect on APM task performances in different thinking modes. The results suggest that a proper think-aloud instruction plays an important role in eliciting non-reactive verbal protocols. When think-aloud is carefully instructed, L2 can be one of the language options in think-aloud instructions. Future research may further explore other factors potentially modulating the reactivity of think-aloud.

References

- Aujla, N., K. Vedhara, M. Walker and N. Sprigg, 2018. Evaluating a stroke-specific version of the illness perception questionnaire-revised, using the Think-Aloud method. *Journal of Health Psychology*: 1-17. Available at: <https://doi.org/10.1177/1359105318781942>.
- Bartłomiejczyk, M., 2006. Strategies of simultaneous interpreting and directionality. *Interpreting*, 8(2): 149-174. Available at: <https://doi.org/10.1075/intp.8.2.03bar>.
- Benson, P., A. Chik, X. Gao, J. Huang and W. Wang, 2009. Qualitative research in language teaching and learning journals, 1997–2006. *The Modern Language Journal*, 93(1): 79-90. Available at: <https://doi.org/10.1111/j.1540-4781.2009.00829.x>.
- Bowles, M.A., 2008. Task type and reactivity of verbal reports in SLA: A first look at a L2 task other than reading. *Studies in Second Language Acquisition*, 30(3): 359-387. Available at: <https://doi.org/10.1017/s0272263108080492>.
- Bowles, M.A., 2010. Concurrent verbal reports in second language acquisition research. *Annual Review of Applied Linguistics*, 30(1): 111-127. Available at: <https://doi.org/10.1017/s0267190510000036>.
- Bowles, M.A. and R.P. Leow, 2005. Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, 27(3): 415-440. Available at: <https://doi.org/10.1017/s0272263105050187>.
- Brandimonte, M.A., G.J. Hitch and D.V.M. Bishop, 1992. Influence of short-term memory codes on visual image processing: Evidence from image transformation tasks. *Journal of Experimental Psychology*, 18(1): 157-165. Available at: <http://dx.doi.org/10.1037/0278-7393.18.1.157>.
- Cohen, A.D., 1984. On taking language tests: What the students report. *Language Testing*, 1(1): 70-81. Available at: <https://doi.org/10.1177/026553228400100106>.
- Cohen, A.D., 1986. Mentalistic measures in reading strategy research: Some recent findings. *English for Specific Purposes*, 5(2): 131-145. Available at: [https://doi.org/10.1016/0889-4906\(86\)90019-0](https://doi.org/10.1016/0889-4906(86)90019-0).
- Ericsson, K.A., 2002. Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: Interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology*, 16(8): 981-987. Available at: <https://doi.org/10.1002/acp.925>.
- Ericsson, K.A. and H.A. Simon, 1980. Verbal reports as data. *Psychological Review*, 87(3): 215-251. Available at: <http://dx.doi.org/10.1037/0033-295X.87.3.215>.
- Ericsson, K.A. and H.A. Simon, 1984/1993. *Protocol analysis: Verbal report as data*. Cambridge, Massachusetts: MIT Press.
- Ericsson, K.A. and H.A. Simon, 1998. How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5(3): 178-186. Available at: https://doi.org/10.1207/s15327884mca0503_3.
- Eveland, J.W.P. and S. Dunwoody, 2000. Examining information processing on the world wide web using think aloud protocols. *Media Psychology*, 2(3): 219-244. Available at: https://doi.org/10.1207/S1532785XMEP0203_2.
- Fallshore, M. and J.W. Schooler, 1995. Verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6): 1608-1623. Available at: <https://doi.org/10.1037/0278-7393.21.6.1608>.
- Flaherty, E.G., 1975. The thinking aloud technique and problem solving ability. *Journal of Educational Research*, 68(6): 223-225. Available at: <https://doi.org/10.1080/00220671.1975.10884753>.
- Guo, C., 2007. *Think aloud protocols*. Beijing: Foreign Language Teaching and Research Press.
- Ivanova, A., 2000. The use of retrospection in research on simultaneous interpreting. In: S. Tirkkonen-Condit, R. Jääskeläinen (Eds.), *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on Empirical Research*. Amsterdam/Philadelphia: John Benjamins Publishing Company. pp: 27-52.
- Kim, H.S., 2002. We talk, therefore we think? A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology*, 83(4): 828-842. Available at: <http://dx.doi.org/10.1037/0022-3514.83.4.828>.
- Leow, R.P. and K. Morgan-Short, 2004. To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26(1): 35-57. Available at: <https://doi.org/10.1017/s0272263104026129>.
- Raven, J.C., J.H. Court and R. J., 1978. *Manual for raven's progressive matrices and vocabulary scales*. London: H. K. Lewis.
- Rossomondo, A.E., 2007. The role of lexical temporal indicators and text interaction format in the incidental acquisition of the Spanish future tense. *Studies in Second Language Acquisition*, 29(1): 39-66. Available at: <https://doi.org/10.1017/s0272263107070027>.
- Russo, J.E., E.J. Johnson and D.L. Stephens, 1989. The validity of verbal protocols. *Memory & Cognition*, 17(6): 759-769. Available at: <https://doi.org/10.3758/BF03202637>.
- Sachs, R. and C. Polio, 2007. Learners' uses of two types of written feedback on an L2 writing revision task. *Studies in Second Language Acquisition*, 29(1): 67-100. Available at: <https://doi.org/10.1017/S0272263107070039>.
- Sachs, R. and B.R. Suh, 2007. Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In: A. Mackey. (Ed.), *Conversational Interaction in Second Language Acquisition: A collection of empirical studies*. Oxford, UK: Oxford University Press. pp: 197-227.
- Sanz, C., H. Lin, B. Lado, H.W. Bowden and C.A. Stafford, 2009. Concurrent verbalizations, pedagogical conditions, and reactivity: Two call studies. *Language Learning*, 59(1): 33-71. Available at: <https://doi.org/10.1111/j.1467-9922.2009.00500.x>.
- Schooler, J.W., S. Ohlsson and K. Brooks, 1993. Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, 122(2): 166-183. Available at: <http://dx.doi.org/10.1037/0096-3445.122.2.166>.
- Schweiger, D.M., 1983. Is the simultaneous verbal protocol a viable method for studying managerial problem solving and decision making? *Academy of Management Journal*, 26(1): 185-192. Available at: <https://doi.org/10.2307/256146>.
- Smagorinsky, P., 1994a. Potential problems and problematic potentials of using talk about writing as data about writing process. In: P. Smagorinsky (Ed.), *Speaking about Writing: Reflections on Research Methodology*. Thousand Oaks, London, New Delhi: Sage Publications. pp: 9-19.
- Smagorinsky, P., 1994b. Think-aloud protocol analysis beyond the black box. In: P. Smagorinsky (Ed.), *Speaking about Writing: Reflections on Research Methodology*. Thousand Oaks, London, New Delhi: Sage Publications. pp: 3-19.
- Smagorinsky, P., 2001. Rethinking protocol analysis from a cultural perspective. *Annual Review of Applied Linguistics*, 21(1): 233-245. Available at: <https://doi.org/10.1017/s0267190501000149>.
- Stinesen, L., 1985. The influence of verbalization on problem-solving. *Scandinavian Journal of Psychology*, 26(1): 342-347. Available at: <https://doi.org/10.1111/j.1467-9450.1985.tb01173.x>.
- Stratman, J.F. and L. Hamp-Lyons, 1994. Reactivity in concurrent think-aloud protocols. In: P. Smagorinsky (Ed.), *Speaking about Writing: Reflections on Research Methodology*. Thousand Oaks, London, New Delhi: Sage Publications. pp: 89-112.

- Wilson, T.D., 1994. The proper protocol: Validity and completeness of verbal reports. *Psychological Science*, 5(5): 249-252. Available at: <https://doi.org/10.1111/j.1467-9280.1994.tb00621.x>.
- Yi'an, W., 1998. What do tests of listening comprehension test?-A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15(1): 21-44. Available at: <https://doi.org/10.1177/026553229801500102>.

Asian Online Journal Publishing Group is not responsible or answerable for any loss, damage or liability, etc. caused in relation to/arising out of the use of the content. Any queries should be directed to the corresponding author of the article.